

LOSSLESS COMPRESSION FOR DVD-AUDIO

JR STUART, PG CRAVEN, MJ LAW

Meridian Audio Ltd, Huntingdon, England
jrs@meridian-audio.com

INTRODUCTION

Meridian Lossless Packing (MLP) is a lossless coding system for high-quality linear PCM audio. This paper gives a brief overview of the DVD-Audio application of MLP. More detail is available in [3] and [2].

For DVD-Audio MLP performs lossless compression of up to 6 channels of up to 24-bit material sampled at rates between 44.1kHz and 192kHz.

Figure 1 gives an overview of the process of losslessly compressing a stream containing multiple audio channels and auxiliary data onto a disc.

1 LOSSLESS COMPRESSION

Lossless coding does not alter the final signal, it ‘packs’ the audio data into a smaller rate and space.

Audio that is of interest to the human listener contains some redundancy. On music signals, the information content varies with time and the input channel information capacity is rarely fully exercised.

The lossless encoder output has a *variable* data rate on normal audio content. Figure 5 illustrates the variation in 30 seconds of 96kHz 24-bit 6-channel orchestral music (original rate 13.824Mbps).

The playing time is determined by the *average data rate* (i.e. compressed file size).

The ARA Proposal [1] described the importance of reducing the *instantaneous peak data rate* for high sampling rates such as 96kHz or 192kHz. MLP tackles this by maximising the compression at all times using this set of techniques:

- Looking for ‘dead air’ – channels that do not exercise all the available word size, *and*
- Channels that do not use the available bandwidth.
- Removing inter-channel correlations.
- Efficiently coding the residual information.
- Smoothing coded information by buffering.

2 HOW DOES IT WORK?

MLP uses some important novel techniques, including:

- Lossless processing.
- Lossless matrixing.
- Lossless use of IIR filters.
- Managed FIFO buffering across transmission.
- Decoder lossless self-check.
- Operation on mixed sample rates.

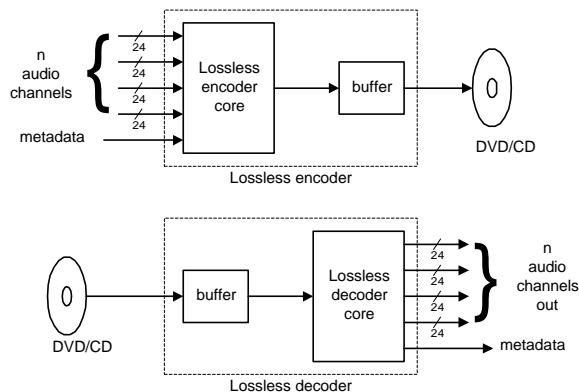


Figure 1: An overview of MLP used on disc.

3 MLP ENCODER

The MLP encoder core is illustrated in Figure 2. The steps for encoding blocks of data are:

1. Incoming channels may be remapped to optimise the use of substreams.
2. Each channel is shifted to recover unused capacity (e.g. less than 24-bit precision or less than full-scale)
3. A lossless matrix technique optimises the channel use by reducing inter-channel correlations.
4. The signal in each channel is de-correlated using a separate predictor for each channel.
5. The de-correlated audio is further optimised using entropy coding.
6. Each substream is buffered using a FIFO memory system to smooth the encoded data rate.
7. Multiple data substreams are interleaved.
8. The stream is packetised for fixed or variable data rate and for the target carrier.

3.1 Lossless Matrix

A multichannel audio mix will usually share some common information between channels.

The MLP encoder uses a special lossless matrix that allows the encoder to reduce correlations, thereby concentrating larger amplitude signals in fewer channels. A simple illustrative example is the tendency of the matrix process to rotate a stereo mix from left/right to sum/difference.

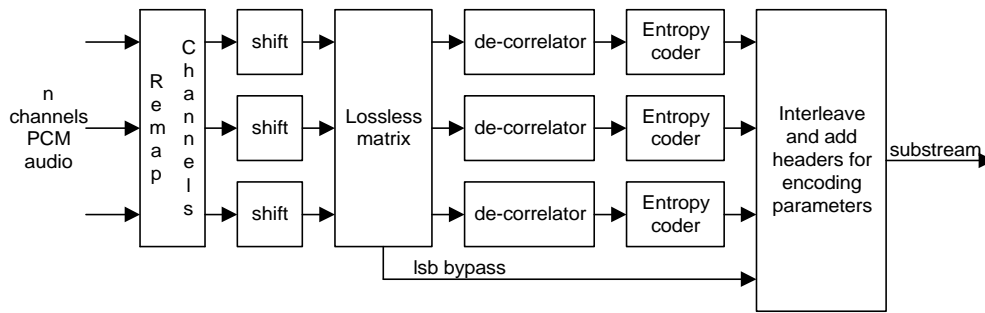


Figure 2: Block diagram of the lossless encoder core.

3.2 Prediction

If the values of future audio samples can be predicted, then it is only necessary to transmit the rules of prediction along with the difference between the estimated and actual signals.

Figure 3 shows the short-term spectrum of a music excerpt. The de-correlator makes its gains by attempting to flatten this spectrum. The Gerzon/Craven theorems show that the level of the optimally de-correlated signal is given by the average of the original signal spectrum when plotted as dB versus linear frequency – see Figure 3. This dB average has less power than the original signal, and the reduction represents the information content of the signal as defined by Shannon.

The MLP encoder uses a separate predictor for each channel. The encoder is free to select IIR or FIR filters up to 8th order from a wide palette. These extensive options ensure that good data reduction can be provided on as many types of audio as possible.

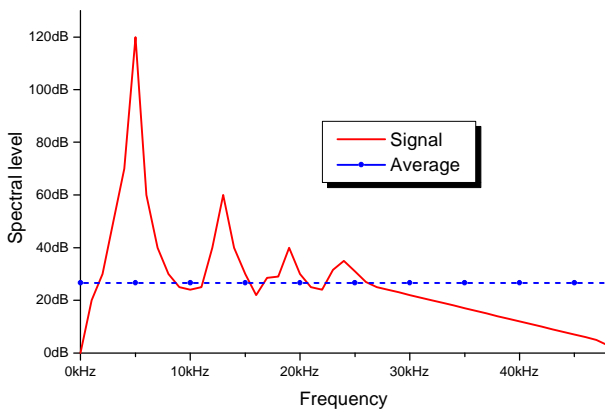


Figure 3: Spectra of a signal and its average level.

3.3 Entropy coding

MLP can select between a number of entropy coding methods to optimally reduce the data following the predictor.

3.4 Buffering

Most audio signals can be predicted, but there will be occasional fragments like sibilants, synthesised noise or percussive events that have high entropy.

MLP uses a particular form of stream buffering that can reduce the variations in transmitted data rate, absorbing transients that are hard to compress.

FIFO memory buffers are used in the encoder and decoder as shown in Figure 4. These buffers give a constant delay across encode and decode. This overall delay is small – typically of the order of 75ms. To allow rapid start-up or cueing, the FIFO management minimises the decoder delay. The decoder buffer is normally empty and is only filled ahead of sections with high instantaneous data rate.

The encoder can use the buffering for a number of purposes, e.g.:

- Keeping the data-rate below a preset (format) limit.
- Minimising the peak data rate over an encoded section. (For an example see Figure 5.)

Figure 6 shows how hard-to-compress signals can be squeezed below a preset format limit. This 30-second 96kHz 24-bit recording features closely recorded cymbals in 6 channels. At the crescendo this signal is virtually random and the underlying compressed data rate is 12.03Mbps. Buffering allows the MLP encoder to hold the transmitted data rate below 9.2Mbps by filling the decoder buffer to a short-term maximum of 86kbyte (bottom curve).

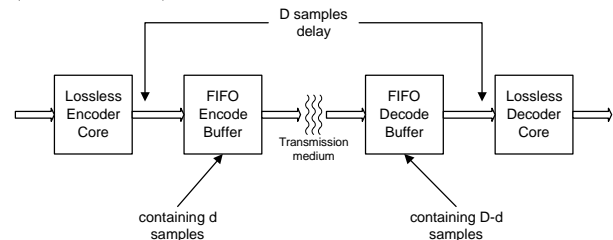


Figure 4: Outlining the buffering used in MLP.

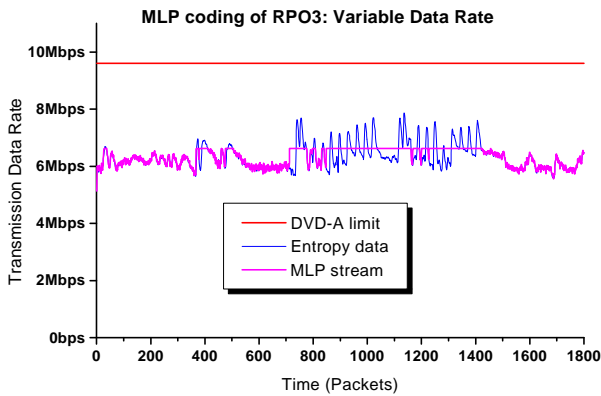


Figure 5: Showing how buffering minimises data rate.

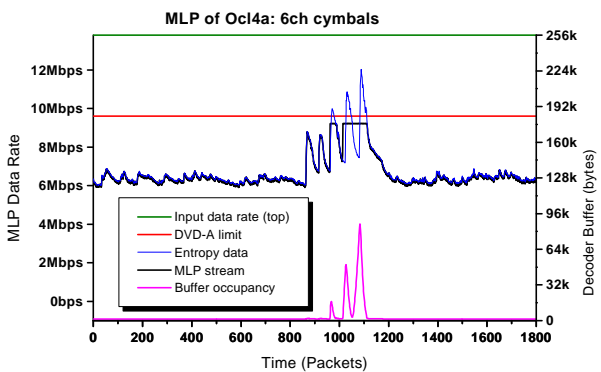


Figure 6: Buffering allows a difficult passage to remain below the format limit.

4 MLP DECODER

The decoder reverses each encoder process. The decoder is relatively low complexity. A decoder for a 2-channel player (capable of extracting a 2-channel stream at 192kHz) requires approximately 27MIPs, while 40MIPs will be required to decode 6 channels at 96kHz.

5 TWO CHANNEL DOWNMIX

It is useful to provide a means for accessing high-resolution multichannel audio streams on 2-channel playback devices. In DVD-Audio, the content provider can place separate multi- and two-channel streams on the disc. However to do this requires separate mix, mastering and authoring processes and uses disc capacity.

In cases where only one multichannel stream is available, then a guided downmix is needed. However, to create such a downmix it is first necessary to decode the full multichannel signal.

5.1 Performing mix-down in the lossless encoder

MLP provides an elegant and unique solution to this problem. The encoder combines lossless matrixing with the use of two substreams in such a way as to optimally encode *both* the Lt/Rt downmix *and* the multichannel version.

This method is shown in Figure 7.

Because the 2-channel downmix is a linear combination of the multichannel mix, then strictly, no new information has been added. Figure 8 shows an example where a downmix is added to a 6-channel segment.

The advantages of this method are considerable:

- The quality of the mix-down is guaranteed. The producer can listen to it at the encoding stage and MLP delivers it bit-accurate to the end user.
- A 2-channel-only player can be very simple.
- A more complex player can access both the 2-channel and multi-channel versions losslessly.
- The downmix coefficients are not fixed, but can be varied with 1ms resolution under artistic control *throughout each track*.

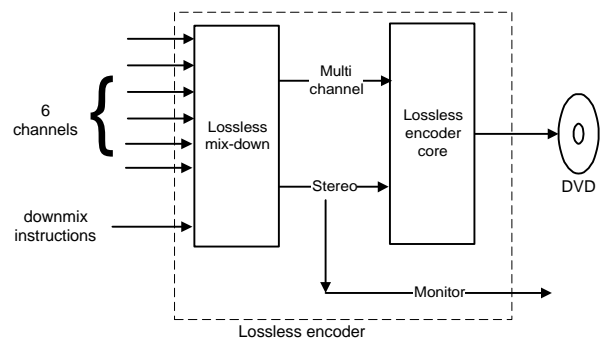


Figure 7: Illustrating encoder downmix.

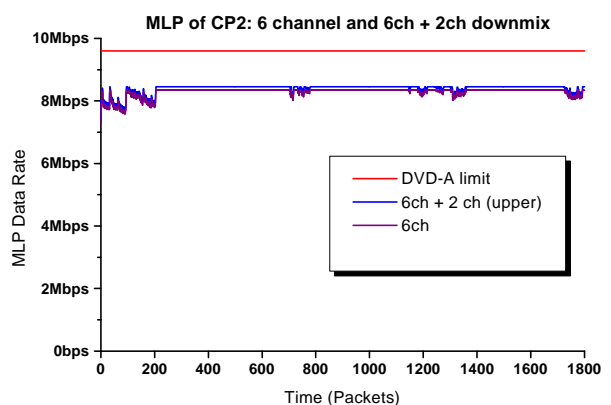


Figure 8: Showing the impact on data rate of adding a 2-channel downmix to 6-channel content.

6 HOW MUCH COMPRESSION?

Table 1 Sampling kHz	Data-rate reduction: bits/sample/channel	
	Peak	Average
48	4	5-11
96	8	9-13
192	9	9-14

The saving of data rate (in bits per original sample) is indicated in Table 1.

In DVD-Audio the peak saving is a key parameter because the encoded stream must always operate below the audio buffer data-rate limit of 9.6Mbps.

DVD-Audio uses a variable-rate MLP stream so the average reduction indicates space saving and therefore the increase in playing time.

6.1 Compression adjustment

A producer may wish to save space used by a recording, or to reduce data rate. Lossless compression extends the number of options.

With MLP, data is automatically saved if the incoming precision is reduced. So, reducing (for example) a few or all channels from 24 to 23 or 22-bit will automatically save data: each bit saves about 8%.

Another option for increasing playing time is to low-pass filter some of the incoming channels. Low-pass filtering reduces the entropy in the signal and the MLP encoder provides a lower data rate. This method is preferable to the use of mixed sample rates because the control is flexible and the stream is homogeneous.

MLP always returns the streams bit-for-bit intact once any mastering adjustments have been made.

6.2 Playing time on DVD-Audio

DVD Audio holds approximately 4.7Gbytes of data and has a maximum data transfer rate of 9.6Mbps.

6-channel 96kHz 24-bit LPCM audio has a data rate of 13.824Mbps, which exceeds this 9.6Mbps limit and would use up the disc capacity in 45 minutes.

So, lossless compression is needed to reduce the data on the disc to extend playing time to the industry norm of 74 minutes *and* to guarantee a *minimum* reduction of 31% in instantaneous data rate.

MLP meets this requirement with a sophisticated encoder, a simple decoder and a specific subset of features limited to two substreams and 6 channels. [2] Here are some typical examples of playing times:

- 5.1 channels 96kHz 24-bit: 100 minutes.
- 6 channels 96kHz 24-bit: 86 minutes.
- 2 channels 96kHz 24-bit: 4 hours.
- 2 channels 192kHz 24-bit: 2 hours.
- 2 channels 44.1kHz 16-bit: 10 hours.
- 1 channel 44.1kHz 16-bit: 20 hours (talking book).

7 SUMMARY

First and foremost MLP is truly lossless and guarantees bit-for-bit delivery of the original data. The decoder can confirm true end-to-end lossless operation.

The main features are:

- Good compression of peak and average data rates.
- Use of both fixed and variable-rate data-streams.
- Automatic savings on bass-effects channels.
- Automatic savings on signals that do not use all of the available bandwidth (e.g. sampled at 96kHz).
- Automatic savings when channels are correlated.
- Hierarchical access to multichannel information.

The bitstream allows robust operation, fast error recovery and rapid cueing (typically recovering in 7ms). The MLP decoder is relatively simple, is hierarchical, and is lossless over different hardware platforms.

In addition to audio, the MLP stream carries additional information of benefit to the decoder, to the content provider and to the end user.

While the highest compression requires sophisticated encoders, near optimal encoding of most music signals can be obtained with much simpler encoders that have modest MIPS requirements and can run in real time with low latency on cheaply available DSP devices. Thus future use in consumer record/playback systems and in radio microphones or other real-time applications is entirely feasible.

8 REFERENCES

- [1] *Acoustic Renaissance for Audio*, 'A Proposal for High-Quality Application of High-Density CD Carriers', private publication (February 1995) Download: <http://www.meridian-audio.com/ara>.
- [2] Toshiba et al., 'DVD Specifications for Read-Only Disc, Part 4: Audio Specification' Version 1.0 (March 1999)
- [3] Gerzon, M.A., Craven, P.G., Stuart, J.R., Law, M. and Wilson, R.J., 'The MLP Lossless Compression System', to be presented at AES 17th International Conference on High Quality Audio Coding, Florence (September 1999).

9 PATENT NOTICE

Several aspects of the MLP encode, decode, packetising and bitstream are the subject of patent application. *Meridian*, *Meridian Lossless Packing* and *MLP* are trademarks of Meridian Audio Ltd.

10 ACKNOWLEDGEMENTS

We are grateful to Tony Faulkner (Green Room Productions) and Warner Music for providing recordings used in the compression examples.